WE CLAIM:

1. A high-throughput method for determining a biochemical function of a protein or polypeptide domain of unknown function comprising:

(A) identifying a putative polypeptide domain that properly folds into a stable polypeptide domain, said stable polypeptide having a defined three dimensional structure;

(B) determining three dimensional structure of the stable polypeptide domain;

(C) comparing the determined three dimensional structure of the stable polypeptide domain to known three-dimensional structures in a protein data bank, wherein said comparison identifies known structures within said protein data bank that are homologous to the determined three dimensional structure; and

(D) correlating a biochemical function corresponding to the identified homologous structure to a biochemical function for the stable polypeptide domain.

2. The method according to claim 1, further comprising the prestep of parsing a target polynucleotide into at least one putative polypeptide domain.

3. The method according to claim 2, wherein said parsing is performed by a first computer algorithm, wherein said first computer algorithm is selected from the group consisting of a computer algorithm capable of determining exon phase boundaries of a polynucleotide, and a computer algorithm capable of determining interdomain boundaries encoded in a polynucleotide.

4. The method of claim 3, further comprising a computer algorithm that compares the putative polypeptide domain sequence with known domain sequences stored within a database.

5. The method of claim 1, wherein said identification of the stable polypeptide domain having a defined three dimensional structure is performed by a set of activity-independent biophysical criteria that assesses the correctness of folding of the polypeptide domain, said set of activity-independent biophysical criteria including at least one of the criteria selected from the group consisting of circular dichroism measurements, ^{1}H-NMR spectroscopy, amide hydrogen-deuterium time course exchange, and thermal denaturation.

6. The method of claim 1, wherein said determination of the three dimensional structure of the stable polypeptide domain is obtained from an NMR spectrometer spectra of said polypeptide domain.

7. The method of claim 6, wherein said NMR spectrometer spectra include one or more spectra selected from the group consisting of nuclear Overhauser effect spectroscopy (NOESY), pulsed-field gradient ^{15}N-heteronuclear single-quantum coherence spectroscopy (PFG-HSQC), pulsed-field gradient triple-resonance HCCNH ^{13}C-^{13}C total correlation spectroscopy (PFG-HCCNH-TOCSY), pulsed-field gradient HCC(CO)NH ^{13}C-^{13}C TOCSY (PFG-HCC(CO)NH-TOCSY), HCCH COSY, HCCNH-TOCSY, HNCO, CANH, CA(CO)NH, CBCNH, CBCA(CO)NH, H(CA)NH, and H(CA)(CO)NH.

8. The method of claim 6, wherein said NMR spectra is analyzed by a second computer algorithm that automatically assigns resonance assignments to the polypeptide sequence.

9. The method of claim 1, wherein said identification of said stable polypeptide domain comprises measuring a time course of amide hydrogen-deuterium exchange.

10. The method of claim 1, wherein prior to step (B), said stable polypeptide domain is optimally solubilized, said optimum solubilization comprising:

i) preparing an array of microdialysis buttons, wherein each of said microdialysis buttons contains at least 1 µl of an approximately 1M solution of said stable polypeptide domain;

ii) dialyzing each member of said array of microdialysis buttons against a different dialysis buffer;

iii) analyzing each of said dialyzed microdialysis buttons to determine whether said stable polypeptide domain has remained soluble; and

iv) selecting the polypeptide domain having optimum solubility characteristics for NMR spectroscopy.

11. The method of claim 1, wherein said comparison of said determined three dimensional structure to said known three-dimensional structures in the protein data bank is performed by a third computer algorithm that is capable of determining 3D structure homology between said determined three dimensional structure and a member of said ~~PDB~~ protein data bank.

a

12. The method according to claim 11, wherein said third computer algorithm is selected from the group consisting of DALI, CATH and VAST.

13. The method of claim 1, wherein said protein data bank is Protein Data Base ("PDB").

14. The method of claim 4, wherein said database contains domain sequence information of known and determined domain sequences.

15. An integrated system for rapid determination of a biochemical function of a protein or protein domain of unknown function:

(A) a first computer algorithm capable of parsing said target polynucleotide into at least one putative domain encoding region;

(B) a designated lab for expressing said putative domain;

(C) an NMR spectrometer for determining individual spin resonances of amino acids of said putative domain;

(D) a data collection device capable of collecting NMR spectral date, wherein said data collection device is operatively coupled to said NMR spectrometer;

(E) at least one computer;

(F) a second computer algorithm capable of assigning individual spin resonances to individual amino acids of a polypeptide;

(G) a third computer algorithm capable of determining tertiary structure of a polypeptide, wherein said polypeptide has had resonances assigned to individual amino acids of said polypeptide;

(H) a database, wherein stored within said database is information about the structure and function of known proteins and determined proteins; and

(I) a fourth computer algorithm capable of determining 3D structure homology between the determined three-dimensional structure of a polypeptide of unknown function to three-dimensional structure of a protein of known function, wherein said protein of known structure is stored within said protein database.

16. The integrated system of claim 15, wherein said fourth computer algorithm is selected from the group consisting of DALI, CATH and VAST.

17. A high-throughput method for determining a biochemical function of a polypeptide of unknown function encoded by a target polynucleotide comprising the steps:

(A) identifying at least one putative polypeptide domain encoding region of the target polynucleotide ("parsing");

(B) expressing said putative polypeptide domain;

(C) determining whether said expressed putative polypeptide domain forms a stable polypeptide domain having a defined three dimensional structure ("trapping");

(D) determining the three dimensional structure of the stable polypeptide domain;

(E) comparing the determined three dimensional structure of the stable polypeptide domain to known three dimensional structures in a Protein Data Bank to determine whether any such known structures are homologous to the determined structure; and

(F) correlating a biochemical function corresponding to the homologous structure to a biochemical function for the stable polypeptide domain.